Author's response to reviews

Title: Evaluating a community based early childhood education and development program in Indonesia: Study protocol for a pragmatic cluster randomized controlled trial with supplementary matched control group.

Authors:

Menno Pradhan (m.p.pradhan@vu.nl)
Sally A Brinkman (sallyb@ichr.uwa.edu.au)
Amanda Beatty (amandabeatty@post.harvard.edu)
Amelia Maika (amelia_maika@yahoo.com)
Elan Satriawan (elansatriawan@gmail.com)
Joppe DeRee (jderee@worldbank.org)
Amer Hasan (ahasan1@worldbank.org)

Version: 2 Date: 19 July 2013

Author's response to reviews: see over
19 July, 2013

Dear Editor,

Please find attached our revised manuscript, “Evaluating a community based early childhood education and development program in Indonesia: Study protocol for a pragmatic cluster randomized controlled trial with supplementary matched control group.” by Pradhan et al.

We would like to thank the reviewer, William Barnett, for his considered and detailed review of our manuscript. Our responses to the reviewer comments are as follows (our responses in italics).

Major Compulsory Revisions

1. Beginning with the first full sentence on page 6 through the next two full sentences, the text is insufficiently clear for readers not already familiar with early childhood services in Indonesia to understand the context for the study. Early childhood services and programs can refer to a wide range of programs not all of which provide education. From this section, the reader cannot know what is meant by "prior to pre-school". Please inform readers when pre-school begins. It is stated that 85% of the poorest children are not participating in any kind of program. However, the reader does not know what is meant by "program." Does this include health programs and informal child care or does it refer only to formal preschool education and care programs? In addition, it is unclear how "poorest" is defined. Does this mean those in poverty, in extreme poverty, or near poverty?

   We have included a description of the range of programs that constitute early childhood programs. Preschool should not have been mentioned. The text should have read “prior to primary school”. We have told the readers when primary school is intended to start. We have clarified the definition of poorest.

2. Several statements made in the text on ages 16 and 17 regarding sample retention and replacement of children from baseline with new children at midline follow-up do not appear to be consistent. At one point it is stated that 101+ children were not available for reinterview because they had moved. Then it is stated in the next paragraph that 101 children had moved and were reinterviewed. Later it is stated that the number of children who moved and were not reinterviewed is unknown. Perhaps this is what the “+” refers to, but it is unclear. Please clarify this matter. Moreover, it is essential that the study clearly reports the number of children who were replaced or that this figure is unknown and the follow-up sample sizes in each cluster if these are not identical to the baseline. Figure 1 does not currently help in this regard as I cannot reconcile the numbers in Figure 1 with the explanation in the text. Although the number of children who died is quite small, it would seem to me that they are properly included in an analysis as participants for whom there are no benefits at follow-up not as participants with unknown status at follow-up.

   There was a totaling error. This has been rectified. 101 children had moved but were tracked down and successfully reinterviewed. 112 children were lost at follow-up. 269 children were added at follow-up. Figure 1
has been updated to reflect not only the observations lost at follow-up but also the additional observations added to make up for them.

3. Statements about the length of delay in testing also appear to be inconsistent. Page 25 states that baseline data collection took place on average 6 months after implementation and midline 7 months after Batch 3 implementation. Page 26 states that effects on baseline child development will be ignored because 3 months of implementation is expected to have little impact. Page 30 states that actual difference in follow-up between baseline and midline was 14 months (I assume this means assessment rather than treatment) which would be consistent with a 3 month initial delay and a 7 month midline delay. The formula for estimating treatment impact implies that delays are equal at baseline and midline. I do not know why the authors would assume that there is no impact of the delay as this is measurable. Comparison of Batch 1 and Batch 3 baseline should provide a suitable measure of the problem caused by the delay. Admittedly this may not be tested with as much power as desired, but it is better than simply assuming there is no difference.

There are two delays. One in data collection and one in implementation. The first is that the baseline was delayed. Thus by the time the baseline was collected Batch 1 villages had been implementing the intervention for 6 months. The second delay occurred when project implementation in batch 3 villages was delayed by 5 months. Thus when the midline was collected 14 months after the baseline, batch 3 villages had been implementing the intervention for 9 months. Thus at baseline, batch 1 villages had been exposed to the intervention for 6 months, while at midline they were exposed for 20 months. Similarly at baseline batch 3 villages had not yet been exposed to the project but by midline, batch 3 villages had been exposed for 9 months. The mention of 7 months was an error.

You are correct in pointing out that a comparison between batch 1 and batch 3 at baseline would yield an estimate of the impact after 6 months of implementation. We have incorporated this into the discussion on analyses.

4. The paper should discuss in the methods sections protections from false positives relating to multiple comparisons. Nowhere does the paper report the number of measures and what will be done to account the number of estimates of treatment effects conducted while taking into account interdependence of the outcome measures.

A discussion on this has been added. We have listed the outcomes we intend to test.

5. The paper should discuss potential heterogeneity of treatment effects and what will be done to investigate these. Based on past research, systematic differences should be expected based on what is actually implemented as a result of the program, which may differ greatly by type as well as quantity from one location to another, and by child and family characteristics. Children who are seriously malnourished at baseline and whose parents have the lowest levels of education might be predicted to benefit more (though the research is not uniform with respect to this issue, there are enough that point in these directions to make this of particular interest).

This discussion has also been added.

Minor Essential Revisions

1. In discussing the analyses to assess validity of the measures the paper omits looking at associations with age, this seems likely to be an unintentional oversight that should be corrected.

Thank you for noting this unintentional oversight. It has been corrected.

Discretionary Revisions

1. Given the nature of the program, the study analysts should consider developing measures of dosage. This will be complicated and could range from a simple measure of amount of funding
actually spent on service to hours of direct services to multiple indicators of the amount of each of a several different services provided.

We will develop measures of dosage.

2. The discussion of the study recognizes that the sample cannot be assumed to be representative because of the way in which it was obtained. However, it would be useful to know how similar or dissimilar it is to Indonesia as a whole with respect to its measured characteristics and what part of the country it might represent. If the sample is sufficiently diverse, it may be possible to simulate results for a more representative sample from the study data. Two key aspects of representativeness might be discussed: local capacity to effectively implement the program; and, the demographics of the children and families served.

We will measure representativeness using demographic characteristics from the rural sub-sample of a nationally representative household survey.

Level of interest: An article of outstanding merit and interest in its field
Quality of written English: Acceptable
Statistical review: No, the manuscript does not need to be seen by a statistician.
Declaration of competing interests:
I declare that I have no competing interests.

We look forward to hearing from you at your earliest convenience. Please address all correspondence to:
Sally Brinkman – sallyb@ichr.uwa.edu.au

Yours sincerely,

Sally Brinkman
Co-Director, Fraser Mustard Centre
Telethon Institute for Child Health Research